02717.P016 *Patent*

APPLICATION FOR UNITED STATES LETTERS PATENT

FOR

FAULT TOLERANT AUTOMATIC PROTECTION SWITCHING FOR DISTRIBUTED ROUTERS

INVENTORS: CEDELL A. ALEXANDER
DONALD B. GROSSER

PREPARED BY:
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN, LLP
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1026
(503) 684-6200

EXPRESS MAIL LABEL NO. EL034437033US

FAULT TOLERANT AUTOMATIC PROTECTIONSWITCHING FOR DISTRIBUTED ROUTERS

FIELD OF THE INVENTION

The invention relates generally to data networking. More specifically, the invention relates to fault tolerant switching within a data network.

BACKGROUND OF THE INVENTION

As the Internet continues to grow exponentially, the bandwidth requirements of the associated networks comprising the Internet similarly continue to grow. The transmission of large quantities of multimedia data, for example, including both audio and video data, has contributed to what has become a seemingly universal demand for increased network bandwidth. In addition to requiring large amounts of bandwidth, real-time multimedia data also tends to be isochronous, in that it is to be received at a destination device at the same rate with which it was transmitted from the sending device. If such real-time multimedia data is delayed during transmission between the source device and the destination device, the resulting image and/or sound quality of the multimedia "stream" may be adversely affected.

When a network segment located between a source device and a destination device becomes unavailable, data en route from the source device to the destination device has to be rerouted so as to bypass the unavailable network segment. Such rerouting of data due to unexpected network outages may add significant delay to the total data transmission time. One cause for unexpected network outages and associated data transmission delays may be attributed to excavation accidents where

one or more network segments are severed. Conventional copper networks, for example, tend to not be very resilient and are quite susceptible to such accidental severing or "backhoe fade."

New technologies are continually being introduced to address the various demands for increased network bandwidth and improved resiliency. Optical networking is one such technology. In general, optical networks utilize optical glass wires or "fibers" to transmit data in the form of light pulses along the fiber. Among other features, optical fiber is capable of carrying much more information than conventional copper wire and is generally not subject to electromagnetic interference as is copper wiring.

Synchronous Optical Network (hereinafter "SONET") is a technology that combines the high-bandwidth network capacity of optical fiber with the resiliency of automatic protection switching. The SONET optical interface is specified in the *American National Standard for Telecommunications - Synchronous optical network (SONET) - Basic description including multiplex structures, rates, and formats, ANSI T1.105-1995*; and SONET Automatic Protection Switching (hereinafter "APS") is specified in the *American National Standard for Telecommunications - Synchronous optical network (SONET) - Automatic Protection Switching, ANSI T1.105.01-1998.* Synchronous Digital Hierarchy (hereinafter "SDH") is the international equivalent of SONET and is standardized by the International Telecommunications Union (ITU). Although SDH may not be specifically referred to herein, the concepts described with respect to SONET nonetheless apply equally to SDH.

The APS protocol describes an architecture in which data signals travelling across one signal path may be automatically switched to another signal path due to a variety of circumstances such as signal degradation or line failure. For example, assume a router is connected to a signal multiplexor by both a working line (i.e. primary circuit) and a protection line (i.e. backup circuit). If the working line were to fail, or signal quality on the working line were to degrade, APS would perform an automatic switchover so that signals would travel to and from the router via the protection line rather than the working line, thereby maintaining communication between the multiplexor and the router.

Several different automatic protection switching schemes are addressed by the APS specification. One such switching scheme is protection switching for linear 1+1 topologies. Linear APS is used to protect tributary SONET lines which connect routers to Add-Drop multiplexors (hereinafter "ADM"), whereas ring APS architectures protect the lines between equipment comprising the SONET ring itself. Figure 1 illustrates an exemplary linear APS 1+1 architecture according to the prior art. As shown in Figure 1, each router is coupled to the ADM and SONET ring by both a working line and a protection line. Accordingly, if either working line were to fail, for example, the corresponding protection line could be used in order to maintain communication with the SONET ring.

Each working line and protection line pair may be coupled to an associated router through either a single SONET interface module or a pair of SONET interface modules. In Figure 1, for example, each working/protection line pair is coupled to an associated router through a pair of interface modules, however, a single module implementation may also be utilized. While the single module implementation protects against line failures, the two module implementation protects against both line and interface module failures. Neither implementation, however, protects against interruptions in network communication due to an entire router failure.

SUMMARY OF THE INVENTION

A fault tolerant switching architecture comprises a working router coupled to a SONET add-drop multiplexor (ADM) through a working line and a protection router coupled to the ADM through a redundant, protection line. Additionally, the working and protection routers are coupled to each other by way of a separate link, referred to as a side-band connection. The working router and protection router comprise a virtual router from the perspective of the neighboring router, which communicates with the virtual router over the SONET network using the Point-to-Point Protocol (PPP). The goal of the architecture is to provide fault tolerance in the event of either a line or router failure. According to one embodiment of the invention, the protection router transmits a heartbeat message to the working router over the side-band connection. If the protection router does not receive a response to the transmitted heartbeat message, the protection router initiates a line switch within the add-drop multiplexor. Once the line switch is complete, the protection router exchanges datagrams with the neighboring router, via the ADM and SONET ring to which the ADM is coupled. The protection router, in so doing, establishes a PPP connection between the protection router and the neighboring router device that is also coupled to the SONET ring. The PPP connection between the protection router and the neighboring router is established with the Link Control Protocol (LCP). The protection router includes a predetermined identifier value that identifies the originator of the request, in the LCP Identifier field of LCP request datagrams. According to the LCP specification, the neighboring router includes the LCP Identifier value received in a request datagram in the corresponding response datagram. The response datagram is transmitted over the SONET ring to the ADM.

Because datagrams received by the ADM from the SONET link are transmitted over both the working and the protect lines, the working router will receive from the ADM the same response that was sent to the protection router. Thus, by examining the identifier field, and recognizing the identifier value as that assigned to the protection router, the working router is able to determine that the line switch to the protection router has occurred. This knowledge enables proper network operation in the case where the protection router loses contact with the working router due to a faulty side-band connection rather than a malfunctioning working router.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is illustrated by way of example, and not by way of limitation in the figures of the accompanying drawings in which like reference numerals refer to similar elements.

Figure 1 is a block diagram illustrating one example of a linear 1+1 APS architecture according to the prior art.

Figure 2 is a block diagram illustrating a prior art, two-router, linear 1+1 APS architecture utilized in one embodiment of the invention.

Figure 3 is a block diagram illustrating a prior art virtual router representation.

Figure 4 illustrates a SONET frame structure including payload and overhead sections.

Figure 5 illustrates formats for a basic HDLC frame and a basic PPP frame.

Figures 6(a)-(c) illustrate the format and possible contents of the *K1* and *K2* bytes.

Figure 7 is a flow diagram illustrating the protection router initiated switching according to one embodiment of the invention.

DETAILED DESCRIPTION

In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the invention. It will be apparent, however, to one skilled in the art that the invention can be practiced without these specific details. In other instances, structures and devices are shown in block diagram form in order to avoid obscuring the invention.

Reference in the specification to "one embodiment" or "an embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the invention. The appearances of the phrase "in one embodiment" in various places in the specification are not necessarily all referring to the same embodiment.

A fault tolerant switching architecture is described herein. More specifically, a novel two-router linear 1+1 APS architecture including both a working and a protection router is described herein. The working router is coupled to an add-drop multiplexor (ADM) through a working line, and the protection router is coupled to the ADM through a protection line. Additionally, the working and protection routers are coupled to each other via a side-band network connection. According to one embodiment of the invention, the protection router transmits a heartbeat message to the working router. If the protection router does not receive a response to the transmitted heartbeat message, the protection router initiates a line switch within the add-drop multiplexor to which both are attached. Once the line switch is complete, the protection router transmits a signal

to the ADM to establish a point-to-point link between the protection router and a neighboring router also coupled to the SONET ring. In one embodiment, the signal is transmitted to the ADM the via protection line, over the SONET ring to the neighboring router, and back to the working router via the line between the ADM and the working router. The signal is an identifier value, which is included in request datagrams and the corresponding response datagrams. The identifier value is used to indicate to the working router that the line switch to the protection router has occurred. Thus, if the protection router loses contact with the working router (i.e., it does not receive a response to the heartbeat message) due to a faulty side-band connection rather than due to a malfunctioning working router, the working router is able to detect that the line switch to the protection router has occurred and can subsequently disable its SONET interface.

The APS standard specifies at least two linear APS 1+1 switching modes: unidirectional protection switching and bidirectional protection switching. In unidirectional protection switching, both head end (i.e. the node executing a bridge) and tail end (i.e. the node that initiates an APS request) monitor the quality of the optical signals on the working line and the protection line such that receiving equipment on either end can independently select the signal with the best quality. In bidirectional protection switching, the switching from one line to another is coordinated. That is, both head and tail ends are synchronized such that both ends select the same signal line. Synchronization is achieved via APS protocols that are carried in the *K1* and *K2* bytes

of the SONET line overhead. The *K1* and *K2* bytes are typically transmitted on the protection line, but may also be transmitted on the working line.

Bidirectional protection switching is advantageous for data communication applications where the working line and the protection line may be terminated in different routers. Figure 2 is a block diagram illustrating a two-router, linear 1+1 APS architecture according to one embodiment of the invention. Referring to Figure 2, a working router and a protection router are shown coupled to SONET ring 20 through ADM 22. In addition to being coupled to ADM 22, the working and protection routers are further coupled to one another by side-band connection 24. Side-band connection 24 represents an out-of-band communications channel that is used to synchronize the working and protection routers' use of the SONET interface. For example, if the working router detects a degradation in signal quality on the working line, it can send a message to the protection router over side-band connection 24. In response, the protection router can then initiate a switch to the protection line so as to maintain communication with SONET ring 20. According to one embodiment of the invention, side-band connection 24 is also used to deliver heartbeat messages that enable the protection router to take over if the working router should fail. In one embodiment, side-band connection 24 is an Ethernet connection that operates out-of-band with respect to SONET ring 20. In other embodiments, however, side-band connection 24 may represent any wired or wireless transmission medium known in the art to transmit heartbeat messages between protection and working routers.

If both the working router and the protection router are configured with the same network address (e.g., IP address), the two routers can appear as a single "virtual" router to other devices on SONET ring 20. Figure 3 is a block diagram illustrating a virtual router representation of the working and protection routers of Figure 2. Virtual router 30 is coupled to neighbor router 33 via Packet over SONET interface module (PoS) 31, SONET ring 20, and PoS 32. PoS 31, as an abstraction of PoS 36 and 28, represents or utilizes the same IP address assigned for PoS 26 and PoS 28 of Figure 2, whereas PoS 32 is configured with a unique IP address. Accordingly, virtual router 30 appears to neighboring router 33 as a single routing device rather than two cooperating routers.

Each PoS represents a network interface that enables a synchronous payload envelope (SPE) to be structured in rows while providing for the transport of any open systems interconnection reference model (OSI-RM) layer 3 packet structure given appropriate framing. The Internet Point-to-Point Protocol (PPP) and OSI-RM High-level Data Link Control Protocol (HDLC) are two data framing protocols used in conjunction with SONET. The PPP definition can be found in the Internet Engineering Task Force's request for comments RFC1661, whereas "PPP in HDLC-like framing" and "PPP over SONET/SDH" are defined in RFC1662 and RFC2615 respectively.

Figure 4 illustrates a SONET frame structure including payload and overhead sections. Each SONET frame is logically organized as a two-dimensional array of bytes with the size of the frame dependent upon the channel rate. The basic SONET channel

consists of frames that contain 810 bytes organized in 9 rows by 90 columns. The first 3 columns are reserved for SONET overhead, whereas the last 87 columns are dedicated to the SPE. A SONET ring provides point-to-point connections between routers within the ring. IP packets therefore utilize a point-to-point protocol, such as PPP or HDLC, to map to a point-to-point link. Figure 5 illustrates formats for a basic HDLC frame and a basic PPP frame with the HDLC framing. Both the HDLC and the PPP frame structures are defined to have minimum and maximum length constraints. Within the HDLC and the PPP frame structures are packets, each of which may vary in size so long as they do not exceed a defined maximum length. In one embodiment of the invention PPP frame structures are used to transport IP packets.

Due to the flexibility of SONET and APS, a variety of line-switching scenarios are possible given the network configuration depicted in Figure 2. For the purposes of this disclosure two of such possible scenarios will be discussed, each of which assumes a bidirectional switching mode implementation. Additionally, each scenario assumes the working router to be active. In the first line switching scenario to be discussed, the working router detects a signal fail or degrade condition on the working line and initiates a line switch to the protection line. In the second line switching scenario to be discussed, the protection router loses contact with the working router and a line switch to the protection line is initiated by the protection router.

SWITCHING SCENARIO 1

In the first switching scenario, it is assumed that the working line is active, and that the working router initiates a switch to the protection line upon detecting a signal fail/degrade condition on the working line. The working router initiates the line switch by transmitting a message to the protection router over side-band connection 24 (of Fig. 2). In one embodiment, communications between the working router and the protection router are conducted via the user datagram protocol (UDP) known in the art. Upon receiving the message transmitted by the working router, the protection router invokes the APS protocol to request that ADM 22 switch to the protection line.

As mentioned previously, the APS protocol is carried in the *K1* and *K2* bytes of the SONET overhead (shown in Fig. 4). Figures 6(a)-(c) illustrate the format and possible contents of the *K1* and *K2* bytes. Referring to Figure 6(a), the first four bits of the *K1* byte represent the switch request bits. These bits provide for the exchange of up to sixteen request messages. Figure 6(b) illustrates ten of the sixteen possible request messages that may be represented by *K1* (bits 1:4). For example, during normal operation, the *K1* request bits are set to "0000" as no request is being made. A signal fail indication, however, would result in the *K1* request bits being set to "1100." The *K1* byte further contains bits five through eight which identify the number of the channel (i.e., working/protection line) carrying the request. For example, in this scenario it is assumed that the working router issues the request, and bits 5-8 of the *K1* byte would therefore be assigned "0001" to reflect this. If the protection router issued the request, bits 5-8 of the *K1* byte would be assigned "0000."

The first four bits of the *K2* byte indicate the channel number that is bridged onto the protection line. If channel "0" is received on bits 5-8 of the *K1* byte, the first four bits of the *K2* byte are set to "0000." Bit five of the *K2* byte indicates the APS architecture implementation, in which a "0" is provisioned for 1+1 architecture. The remaining three bits of the *K2* byte indicate various modes or status messages available as shown in Figure 4(c).

	Protect Router → ADM		**ADM → Protect Router**	
	K1 Byte	***K2 Byte***	***K1 Byte***	***K2 Byte***
(a)	0000 0000	0000 0 101	0000 0000	0000 0 101
(b)	1010 0001	0000 0 101	0000 0000	0000 0 101
(c)	1010 0001	0000 0 101	0010 0001	0001 0 101
(d)	1010 0001	0001 0 101	0010 0001	0001 0 101
(e)	1010 0001	0001 0 101	0010 0001	0001 0 101

TABLE 1

Table 1 (above) illustrates one example of an APS protocol exchange between the protection router and ADM 22 (shown in Fig. 2), for switching from the working line to the protection line. In row (a), the first four bits of *K1* indicate that no request is being made, and the last four bits indicate that the channel number is set to the protection line by default. Similarly, the *K2* byte is in a steady state, with the last three bits indicating that a bidirectional switching mode is active. In row (b), the protection router receives a signal degrade message from the working router over side-band connection 24, and sends a signal degrade request for channel 1 (i.e., working line) to ADM 22. In row (c),

ADM 22 acknowledges the signal degrade request by sending a reverse request for channel 1 in the *K1* byte. The *K2* byte indicates that ADM 22 has bridged channel 1 to the protection line. In row (d), the protection router selects (i.e., receives) channel 1 data from the protection line based on the received *K2* byte, and indicates that channel 1 is bridged to the protection line via *K2*. Lastly in row (e), ADM 22 selects (i.e., receives) channel 1 data from the protection line based upon the received *K2* byte.

Once the APS line switch indicated in Table 1 has completed, the protection router sends a message to the working router over side-band connection 24 indicating that the line switch has been performed. The working router then responds by taking down the SONET interface module and initiating a routing topology update. Similarly, the protection router brings its SONET interface module up, and advertises availability of routes accessible via its SONET interface module. Thus, although the switch to the protection line is performed by *K1* and *K2* bytes carried on the protection line, the line switch is actually initiated by the working router.

SWITCHING SCENARIO 2

In the second switching scenario, the working line is again assumed to be active, however, the switch to the protection line is initiated by the protection router rather than the working router. In one embodiment, the protection router initiates the line switch based at least in part upon the protection router losing contact with the working router. In one embodiment, the protection router periodically transmits heartbeat messages to the working router over side-band connection 24 (shown in Fig. 2). In one embodiment,

side-band connection 24 represents an Ethernet-based connection between the working and protection routers. If the protection router detects that a loss of contact with the working router has occurred, for example, by not receiving a response to the transmitted heartbeat messages, the protection router, without additional information, will not be able to discern whether the communication lapse is due to a failure with the working router, or due to a failure of the side-band (e.g., Ethernet) connection, for example. In such an event, according to one embodiment of the invention, an identifier value is included within a PPP configuration frame sent from the protection router, via the ADM and SONET ring, in order to notify the working router that a line switch has occurred. In one embodiment of the invention, the protection router initiates the line switch by invoking the APS protocol as is illustrated above in Table 1 with respect to switching scenario 1. After the switch has been performed, the protection router brings its SONET interface module up and advertises its availability via one or more routing protocols known in the art.

As part of bringing its SONET interface module up, the protection router initializes a PPP link between the protection router and the neighboring router on the SONET ring. In addition to specifying the encapsulation method, PPP also specifies the Link Control Protocol (LCP) which is described more fully in RFC1661. According to RFC1661, if the PPP protocol field (shown in Fig. 5) contains "c021h", exactly one LCP packet will be encapsulated within the PPP information field. All LCP-compliant packets include an 8-bit code field that identifies the LCP packet type, and an 8-bit identifier field that is used as an aid in matching LCP requests and replies. In one embodiment of the

invention, the identifier field of the LCP packet is used to notify the working router that a line switch has occurred. This notification necessarily occurs via the ADM and SONET ring. Advantageously, the neighboring router, a peer in the PPP link with the protection router, requires no modification for the described embodiment to operate.

Figure 7 is a flow diagram illustrating one embodiment of the protection router initiated switching operation described above. Both the working line and the working router are assumed to be up and active and operating in a steady state where no requests are being made (72). According to the SONET APS standard, *K1* and *K2* overhead bytes are transmitted on the protect line, and payload data is transmitted on the working line. Because the working router may not receive the content of the *K1* and *K2* bytes, the working router is not aware of which switch line is active.

Periodically, the protection router transmits heartbeat messages to the working router over the previously described side-band connection (73). Under normal operating conditions, the working router would respond to the heartbeat messages within a predetermined amount of time. Therefore, if a response to one or more of the heartbeat messages is not received by the protection router within the allotted amount of time, the protection router infers that a communication problem exists (74). Because the problem may be caused by failure of the working router or by failure of the side-band connection, the protection router initiates a switch from the working line to the protection line (75), and sends an LCP configure request via the ADM (76) to the peer PPP router, that is, the neighboring router on the SONET ring. According to one

embodiment of the invention, the protection router includes within the LCP, a predetermined identifier value that identifies the originator of the request. In one embodiment of the invention, the most significant bit of the LCP identifier field is used to indicate the originator of the LCP request. For example, a "0" in the most significant bit location of the LCP identifier field signifies that the LCP request was originated by the working router, and a "1" in the most significant bit location of the LCP identifier field signifies that the LCP request was originated by the protection router.

Once the protection router transmits the LCP configure request including the identifier value via the ADM (76), the request is received by a neighboring SONET device, such as a router. Because the neighbor device, for example, believes it is communicating with a single "virtual" router rather than two separate routers, the neighbor device interprets the configure request as a renegotiation and responds to the request with an LCP configure response (e.g., a configure-ack datagram) including the identifier selected by the originator (i.e., protection router) (77). The response is transmitted over the SONET ring to the ADM. Because LCP packets received at the ADM are then transmitted from the ADM over both the working and the protect lines, the working router will receive from the ADM the same response that was sent to the protection router (78). Thus, by examining the identifier field, and recognizing the identifier value as that assigned to the protection router, the working router is able to determine that the protection router is in control of the PPP link (79). For example, according to one embodiment of the invention, whenever the working router receives an LCP packet that contains an identifier field value having the most significant bit set to

"1", the working router recognizes that the protection router is in control of the PPP link. Thereafter, the working router disables its SONET interface so as to not advertise inaccurate routing information.

Additionally, the working router may periodically transmit configure request datagrams containing an identifier value with the most significant bit set to "0", in order to detect a condition where the protection router has failed and the ADM has performed a line switch back to the working line. By inspecting each configure response received on the working line, the working router is able to recognize this condition when the most significant bit of a received identifier value is set to "0". Upon recognizing this condition, the working router will activate its SONET interface.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes can be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.